AD_____

Award Number: W81XWH-10-1-0675

TITLE: Genome-Wide Association Study in African-Americans with Systemic Lupus Erythematosus

PRINCIPAL INVESTIGATOR: John Harley, M.D., Ph.D.

CONTRACTING ORGANIZATION: Children's Hospital Medical Center
Cincinnati, Ohio, 45229-3026

REPORT DATE: September 2013

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED**
September 2013	Annual	1 September 2012 – 31 August 2013

4. TITLE AND SUBTITLE

Genome-Wide Association Study in African-Americans with Systemic Lupus Erythematosus

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-10-1-0675

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

John Harley, M.D., Ph.D.

E-Mail: John.Harley@cchmc.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Children's Hospital Medical Center
Cincinnati, Ohio, 45229-3026

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Systemic lupus erythematosus (lupus) is a potentially deadly systemic autoimmune disease that disproportionately afflicts women and African-Americans. This project is designed to discover genes that increase the risk of lupus in African-American. Our goal was to expand the genotyping density by genotyping a subset of our African-American lupus cases and controls on the OMNI-1S platform and then to exploit this genotyping with sequencing and follow up genotyping in an effort to identify the genes that alter disease risk. Two years ago we abandoned our hope of genotyping the control samples from Detroit and instead found and have already genotyped 3000 African-American controls on the OMNI-Express. (These reagents were purchased by our collaborator.) At this point the genotyping is completed and we are working on the quality control and data analysis. As we work through the remaining issues (population stratification, imputation, and detailed exploration of positive findings) we are hopeful to be composing our genome wide association study manuscript in the months to come and then turn our attention to the follow up studies with sequencing data and follow up genotyping. The project now appears to be spectacularly productive with a substantial number of previously unknown genetic associations that are newly detected by this project.

15. SUBJECT TERMS
Lupus, African-American, Genome Wide Association Study (GWAS), genetic association, genetic susceptibility, candidate gene association study, genotyping

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					USAMRMC
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(include area code)*
U	U	U	UU	8	

Table of Contents

1. Abstract/Introduction (SF298 requirement)

Systemic lupus erythematosus (lupus) is a potentially deadly systemic autoimmune disease that disproportionately afflicts women and African-Americans. This project is designed to discover genes that increase the risk of lupus in African-American. Our goal was to expand the genotyping density by genotyping a subset of our African-American lupus cases and controls on the OMNI-1S platform and then to exploit this genotyping with sequencing and follow up genotyping in an effort to identify the genes that alter disease risk. Two years ago we abandoned our hope of genotyping the control samples from Detroit and instead found and have already genotyped 3000 African-American controls on the OMNI-Express. (These reagents were purchased by our collaborator.) At this point the genotyping is completed and we are working on the quality control and data analysis. As we work through the remaining issues (population stratification, imputation, and detailed exploration of positive findings) we are hopeful to be composing our genome wide association study manuscript in the months to come and then turn our attention to the follow up studies with sequencing data and follow up genotyping. The project now appears to be spectacularly productive with a substantial number of previously unknown genetic associations that are newly detected by this project.

2. Body

Since the population history of African ancestry appears to reach back in time much further to the most recent small founder population (~200,000 years) than the other major human ancestries (<50,0000 years for Asian, European, or Amerindian), the extent of linkage disequilibrium is much lower in population samples of African ancestry. This means that the usual approach for finding genetic association using haplotype block tagged markers will be less successful in this ancestry. One way to partially compensate for this problem is to increase marker density, which is what we are funded to do in this project in our genetic study of systemic lupus erythematosus (lupus).

Lupus in African-Americans is more severe and more deadly than in other populations, and especially so compared to European-Americans. Indeed, lupus afflicts women ten times more frequently than men with a strong tendency to strike during the child-bearing years and is relatively common among the Active Duty Military.

3. Key Research Accomplishments

- Our DOD project is a component of a larger project to more fully characterize African-American genetic association with lupus by genome wide association. Genome wide association genotyping is completed. We continue the intensive evaluation of the data, removing markers and samples that appear to have the potential of artifact from stratification, poor clustering, batch effects, reduced marker calling, and disproportions in controls. At this point we are working with data from all sources from 7,300 DNA samples. Of these this project contributed 574 controls and 1,590 cases evaluated on the OMNI-1S platform. These results are complemented by genotyping from 434 controls and 2,359 cases on the OMNI-1 platform. (Together, the OMNI-1S and OMNI-1 are almost equivalent to the OMNI-2.5.) We have 1,536 controls available with genotyping from the OMNI-2.5 platform, available from dbGaP. In addition, we genotyped 3,985 African-American controls on the OMNI-Express platform in a collaboration with Mt. Sinai in New York City. The merged data provide nearly 3 million single nucleotide markers from 7,300 subjects for a dataset that will have the power for inquiry equivalent to nearly 22 billion genotypes when the imputation across the platforms is completed.

- We have discovered presumed artifacts that originate from batch effects between controls sets, which we have eliminated from the data.

- The major issue has been that the imputation of genotypes against the 1000 genomes project was technically difficult to perform and has required a year of diligent effort to solve. As a consequence we have an abundance of positive associations that do not appear to be robust. Consequently, we submitted a request for extension without additional funds in order to address the issues raised by these results, which was submitted to the DoD on 27 August, 2013. We are awaiting a decision on this request. The following four bullets are taken directly from that request.

- The delays in our plans came from the genotyping data having been performed on four different array platforms. This made the imputation of genotyping very difficult, crashing our computer system repeatedly and requiring much sophisticated informatic intervention and experimentation. Indeed, we started the imputation in September 2012 and did not finish the somatic chromosomes until July 2013. The X-chromosome is now being imputed and we hope will be completed before the end of September.

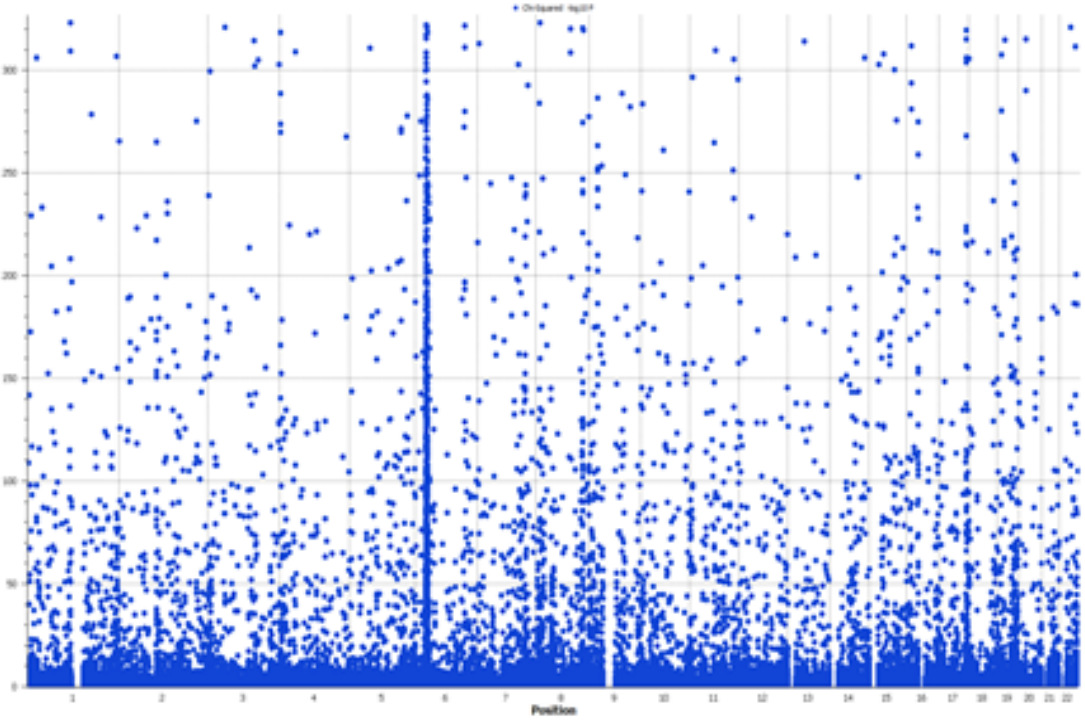

Figure 1. The –log (p) of association is plotted (0 to 340) against the genome position (somatic chromosome number) for the genotyped and imputed results before data cleaning. Our suspicion is that many of these purported associations are spurious and that only a subset are robust. The experiments proposed will attempt to identify such robust associations.

- We cannot order the confirmatory reagents until we are settled on what results need our experimental attention. At this point, we are close to knowing what results need closer

experimental scrutiny, but we are not quite ready, given the still pending X chromosome results. We have saved some funding for this purpose and would intend to do this and any sequencing that we might be able to afford from the remaining funds.

- The outstanding issue has to do with results that are imputed. With a small error accompanies imputation, which is compounded when there is any issue with the genotyping data upon which the imputation is based. Consequently, we approach the imputation results with some reservation that they are correct. Since we have so many apparently positive results we evaluated these data to find the most trustworthy results that we would then intend to publish. Figure 1 presents the initial imputation results for the somatic chromosomes. We imposed additional data quality criteria on these results requiring that there be a 90% call rate, that Hardy Weinberg proportions have $p>0.001$ in the controls, that the samples done with OMNI-1 controls did not differ from the other controls by $p<0.001$, that there be a genotyped marker with an association of $p>0.001$ that is within 20 kb that is in disequilibrium with an $r^2>0.2$. After the loci that do not pass these criteria are removed, we are left with the HLA region, which contains multiple imputed associations, and 100 others, as presented in Figure 2, that have a predicted genome wide imputed association ($p<5x10^{-8}$). These results are extraordinary and we think it important to determine how many of these are robust, which we would test by direct genotyping in a subset of the subjects.

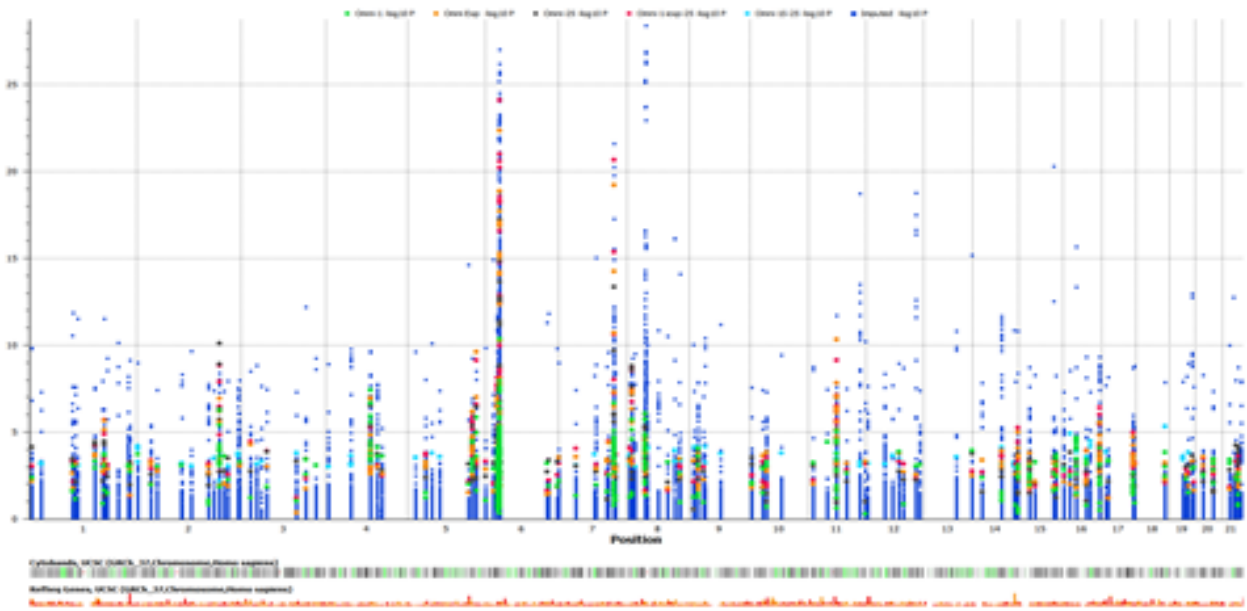

Figure 2. The –log (p) of association (0 to 29) is plotted against the genome position (somatic chromosome number) for the genotyped and imputed results after data cleaning, leaving the HLA region on Chromosome 6 and 101 other genome wide significant associations. Blue dots are imputed genotypes; each the other colors indicate genotyped loci from one of the three contributing arrays.

- The critical experiment that we hope to perform with the extension without additional funds (EWAF) is the actual genotyping of as many of the imputed associations as possible in as many subjects as we can afford. Consequently, we will use our remaining funds to perform this critical experiment. Perhaps, we will be able to establish the validity of the association of at least some of these loci (Figure 2) with actual experimental data.

- Studies published this past year have presented novel findings in a variety of settings. Molinaros et al (1) shows how applying the strategy of admixture mapping can lead to the detection of genetic effects, in this case showing the complex relationship of IFIH1 to lupus. Vaughn et al (2) reviewed the current state of lupus genetics with an emphasis on B cell signaling. Namjou et al (3) discovered a unique variant of the C1qA gene, known in the null form or when anti-C1qA autoantibodies are present to be associated with lupus. Finally, Kim et al published data showing that the ICAM1-OCAM4-ICAM5 locus is associated with lupus.

- We look forward to completing this project and hope that we will have a comparatively productive final report next year.

4. Reportable Outcomes

1. Molineros JE, Maiti AK, Sun C, Looger LL, Han S, Kim-Howard X, Glenn S, Adler A, Kelly JA, Niewold TB, Gilkeson GS, Brown EE, Alarcon GS, Edberg JC, Petri M, Ramsey-Goldman R, Reveille JD, Vila LM, Freedman BI, Tsao BP, Criswell LA, Jacob CO, Moore JH, Vyse TJ, Langefeld CL, Guthridge JM, Gaffney PM, Moser KL, Scofield RH, Alarcon-Riquelme ME, Network B, Williams SM, Merrill JT, James JA, Kaufman KM, Kimberly RP, Harley JB, and Nath SK. (2013) Admixture mapping in lupus identifies multiple functional variants within IFIH1 associated with apoptosis, inflammation, and autoantibody production, *PLoS Genet 9*, e1003222. PMCID: PMC3575474.

2. Vaughn SE, Kottyan LC, Munroe ME, and Harley JB. (2012) Genetic susceptibility to lupus: the biological basis of genetic risk found in B cell signaling pathways, *Journal of leukocyte biology 92*, 577-591. PMC Journal -In Process

3. Namjou B, Keddache M, Fletcher D, Dillon S, Kottyan L, Wiley G, Gaffney PM, Wakeland BE, Liang C, Wakeland EK, Scofield RH, Kaufman K, and Harley JB. (2012) Identification of novel coding mutation in C1qA gene in an African-American pedigree with lupus and C1q deficiency, *Lupus 21*, 1113-1118. PMCID: PMC3508769.

4. Kim K, Brown EE, Choi CB, Alarcon-Riquelme ME, Biolupus, Kelly JA, Glenn SB, Ojwang JO, Adler A, Lee HS, Boackle SA, Criswell LA, Alarcon GS, Edberg JC, Stevens AM, Jacob CO, Gilkeson GS, Kamen DL, Tsao BP, Anaya JM, Guthridge JM, Nath SK, Richardson B, Sawalha AH, Kang YM, Shim SC, Suh CH, Lee SK, Kim CS, Merrill JT, Petri M, Ramsey-Goldman R, Vila LM, Niewold TB, Martin J, Pons-Estel BA, Genles, Vyse TJ, Freedman BI, Moser KL, Gaffney PM, Williams A, Comeau M, Reveille JD, James JA, Scofield RH, Langefeld CD, Kaufman KM, Harley JB, Kang C, Kimberly RP, and Bae SC. (2012) Variation in the ICAM1-ICAM4-ICAM5 locus is associated with systemic lupus erythematosus susceptibility in multiple ancestries, *Ann Rheum Dis 71*, 1809-1814. PMCID: PMC3466387.

5. Conclusions

The planned project to expand the density and number of markers in a case-control study of African-American systemic lupus erythematosus is proceeding toward completion.

6. References

1. Molineros JE, Maiti AK, Sun C, Looger LL, Han S, Kim-Howard X, Glenn S, Adler A, Kelly JA, Niewold TB, Gilkeson GS, Brown EE, Alarcon GS, Edberg JC, Petri M, Ramsey-Goldman R, Reveille JD, Vila LM, Freedman BI, Tsao BP, Criswell LA, Jacob CO, Moore JH, Vyse TJ, Langefeld CL, Guthridge JM, Gaffney PM, Moser KL, Scofield RH, Alarcon-Riquelme ME, Network B, Williams SM, Merrill JT, James JA, Kaufman KM, Kimberly RP, Harley JB, and Nath SK. (2013) Admixture mapping in lupus identifies multiple functional variants within IFIH1 associated with apoptosis, inflammation, and autoantibody production, *PLoS Genet 9*, e1003222. PMCID: PMC3575474.

2. Vaughn SE, Kottyan LC, Munroe ME, and Harley JB. (2012) Genetic susceptibility to lupus: the biological basis of genetic risk found in B cell signaling pathways, *Journal of leukocyte biology 92*, 577-591. PMC Journal -In Process

3. Namjou B, Keddache M, Fletcher D, Dillon S, Kottyan L, Wiley G, Gaffney PM, Wakeland BE, Liang C, Wakeland EK, Scofield RH, Kaufman K, and Harley JB. (2012) Identification of novel coding mutation in C1qA gene in an African-American pedigree with lupus and C1q deficiency, *Lupus 21*, 1113-1118. PMCID: PMC3508769.

4. Kim K, Brown EE, Choi CB, Alarcon-Riquelme ME, Biolupus, Kelly JA, Glenn SB, Ojwang JO, Adler A, Lee HS, Boackle SA, Criswell LA, Alarcon GS, Edberg JC, Stevens AM, Jacob CO, Gilkeson GS, Kamen DL, Tsao BP, Anaya JM, Guthridge JM, Nath SK, Richardson B, Sawalha AH, Kang YM, Shim SC, Suh CH, Lee SK, Kim CS, Merrill JT, Petri M, Ramsey-Goldman R, Vila LM, Niewold TB, Martin J, Pons-Estel BA, Genles, Vyse TJ, Freedman BI, Moser KL, Gaffney PM, Williams A, Comeau M, Reveille JD, James JA, Scofield RH, Langefeld CD, Kaufman KM, Harley JB, Kang C, Kimberly RP, and Bae SC. (2012) Variation in the ICAM1-ICAM4-ICAM5 locus is associated with systemic lupus erythematosus susceptibility in multiple ancestries, *Ann Rheum Dis 71*, 1809-1814. PMCID: PMC3466387.

7. Appendices

NONE.